Fast interior point solvers for H1-regularized PDE-constrained optimization problems

Citation for published version:
Pearson, J & Gondzio, J 2016, Fast interior point solvers for H^1-regularized PDE-constrained optimization problems. in V Bach & H Fassbender (eds), *Joint 87th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM) and Deutsche Mathematiker-Vereinigung.* vol. 16, Wiley-VCH, pp. 737-738. https://doi.org/10.1002/pamm.201610357

Digital Object Identifier (DOI):
10.1002/pamm.201610357

Published In:
Joint 87th Annual Meeting of the International Association of Applied Mathematics and Mechanics (GAMM) and Deutsche Mathematiker-Vereinigung

Fast interior point solvers for H^1-regularized PDE-constrained optimization problems

John W. Pearson[1,*] and **Jacek Gondzio**[2]

[1] School of Mathematics, Statistics and Actuarial Science, University of Kent, Cornwallis Building (East), Canterbury, CT2 7NF, UK
[2] School of Mathematics, University of Edinburgh, James Clerk Maxwell Building, The King's Buildings, Peter Guthrie Tait Road, Edinburgh, EH9 3FD, UK

We consider Newton systems arising from the interior point solution of PDE-constrained optimization problems. In particular, we examine problems where the control variable is regularized by an H^1-norm within the cost functional. We present preconditioned iterative methods for the resulting matrix systems, and justify the potency of our approach through numerical experiments.

1 Problem statement

We consider linear, time-independent PDE-constrained optimization problems with additional bound constraints, of the form:

$$\begin{aligned}
\min_{y,u} \quad & \frac{1}{2}\|y-\widehat{y}\|^2_{L_2(\Omega)} + \frac{\beta}{2}\|u\|^2_{H^1(\Omega)} \\
\text{s.t.} \quad & \mathcal{D}y = u, \quad \text{in } \Omega, \\
& y = f, \quad \text{on } \partial\Omega, \\
& y_a \leq y \leq y_b, \quad \text{a.e. in } \Omega, \\
& u_a \leq u \leq u_b, \quad \text{a.e. in } \Omega,
\end{aligned}$$

where y and u denote state and control variables which we wish to determine, $\widehat{y}$ is a given desired state, $\beta > 0$ a regularization parameter, $\mathcal{D}$ some differential operator, and y_a, y_b, u_a, u_b prescribed bound constraints on the state and control variables. The problem is solved in a domain $\Omega \subset \mathbb{R}^d$, $d \in \{2,3\}$, with boundary $\partial\Omega$. We highlight that the regularization term corresponding to the control is $\|u\|^2_{H^1(\Omega)} = \|u\|^2_{L_2(\Omega)} + \|\nabla u\|^2_{L_2(\Omega)}$ – problems with an L_2-norm regularization term for the control are considered in [2], as are time-dependent variants of such problems.

Applying an interior point method and a *discretize-then-optimize* approach, as in [2], leads to a discrete Lagrangian of the following form:

$$\begin{aligned}
\mathcal{L}(\vec{y},\vec{u},\vec{p}) = {} & \frac{1}{2}\vec{y}^\top M\vec{y} - \vec{y}_d^\top\vec{y} + \frac{\beta}{2}\vec{u}^\top(M+K)\vec{u} + \vec{p}^\top\left(A\vec{y} - M\vec{u} - \vec{f}\right) \\
& - \mu\sum_j \log\left(y_j - y_{a,j}\right) - \mu\sum_j \log\left(y_{b,j} - y_j\right) - \mu\sum_j \log\left(u_j - u_{a,j}\right) - \mu\sum_j \log\left(u_{b,j} - u_j\right),
\end{aligned}$$

of which we wish to find the stationary point. Here $\vec{p}$ and $\vec{y}_d$ correspond to the discretized adjoint variable and desired state, y_j, $y_{a,j}$, $y_{b,j}$, u_j, $u_{a,j}$ and $u_{b,j}$ represent the values of y, y_a, y_b, u, u_a and u_b at the j-th finite element node, K and M are finite element stiffness and mass matrices, A is the finite element matrix related to the PDE operator $\mathcal{D}$, and μ is the chosen barrier parameter within the interior point method.

Applying Newton iteration (with Newton steps $\vec{s}_y$, $\vec{s}_u$, $\vec{s}_p$, and previous iterates $\vec{y}^*$, $\vec{u}^*$, $\vec{p}^*$) to the resulting first-order optimality conditions leads to matrix systems of the form

$$\left[\begin{array}{ccc} M+D_y & 0 & A^\top \\ 0 & \beta(M+K)+D_u & -M \\ A & -M & 0 \end{array}\right]\left[\begin{array}{c} \vec{s}_y \\ \vec{s}_u \\ \vec{s}_p \end{array}\right] = \left[\begin{array}{c} \mu(Y-Y_a)^{-1}\vec{e} - \mu(Y_b-Y)^{-1}\vec{e} + \vec{y}_d - M\vec{y}^* - A^\top\vec{p}^* \\ \mu(U-U_a)^{-1}\vec{e} - \mu(U_b-U)^{-1}\vec{e} - \beta(M+K)\vec{u}^* + M\vec{p}^* \\ \vec{f} - A\vec{y}^* + M\vec{u}^* \end{array}\right],$$

where

$$D_y = (Y-Y_a)^{-1}Z_{y,a} + (Y_b-Y)^{-1}Z_{y,b}, \qquad D_u = (U-U_a)^{-1}Z_{u,a} + (U_b-U)^{-1}Z_{u,b}.$$

Here, Y, U, Y_a, Y_b, U_a, U_b are diagonal matrices containing the entries of y, u (from the previous Newton iteration), y_a, y_b, u_a, u_b at each finite element node, and $Z_{y,a}$, $Z_{y,b}$, $Z_{u,a}$, $Z_{u,b}$ contain entries of the form $\frac{\mu}{y-y_a}$, $\frac{\mu}{y_b-y}$, $\frac{\mu}{u-u_a}$, $\frac{\mu}{u_b-u}$, respectively.

* Corresponding author: e-mail j.w.pearson@kent.ac.uk,

2 Preconditioning

We now note that the Newton system is of saddle point form (see [1] for a survey of such systems). Using the justification provided in [3, 4], we consider the block triangular preconditioners

$$\mathcal{P}_1 = \begin{bmatrix} M + D_y & 0 & 0 \\ 0 & \beta(M+K) + D_u & 0 \\ A & -M & -\widehat{S}_1 \end{bmatrix}, \qquad \mathcal{P}_2 = \begin{bmatrix} -\widehat{S}_2 & 0 & A^\top \\ 0 & \beta(M+K) + D_u & -M \\ 0 & -M & 0 \end{bmatrix},$$

where $\widehat{S}_1$ and $\widehat{S}_2$ are derived using 'matching strategies', as follows:

$$\widehat{S}_1 = \left(A + \widehat{M}\right)(M + D_y)^{-1}\left(A + \widehat{M}\right)^\top,$$
$$\widehat{S}_2 = -\left(A + \widehat{M}\right)^\top M^{-1}(\beta(M+K) + D_u)M^{-1}\left(A + \widehat{M}\right),$$

with $\widehat{M} = M\left[\mathrm{diag}\left(\beta(M+K) + D_u\right)\right]^{-1/2}\left[\mathrm{diag}(M + D_y)\right]^{1/2}$. Eigenvalue analysis concerning similar preconditioners for interior point methods can be found in [2].

In practice it is sensible to use multigrid methods to apply the inverse of the $(1,1)$-block, and the approximate Schur complement inverses $\widehat{S}_1^{-1}$ and $\widehat{S}_2^{-1}$. We apply both preconditioners $\mathcal{P}_1$ and $\mathcal{P}_2$ within the GMRES algorithm.

3 Numerical results

We now implement an interior point method, coupled with our GMRES solver (with preconditioners $\mathcal{P}_1$ and $\mathcal{P}_2$), for a particular test problem. In more detail, we set $\mathcal{D} = -\nabla^2$, $\widehat{y} = e^{-64\left((x_1-0.5)^2+(x_2-0.5)^2\right)}$, where $\mathbf{x} = [x_1, x_2]^\top \in \Omega = [0,1]^2$, with state and control constraints prescribed based on the physical properties of the problem. The iterative solvers are run to a tolerance of 10^{-8}, with the outer (interior point) solver set to a tolerance of 10^{-6}. We solve for a range of step-sizes h and values of β, using MATLAB R2015a, on a quad-core 3.2 GHz processor. We observe that the number of interior point iterations, as well as the GMRES iteration count, behave robustly for a wide range of parameters. We therefore conclude that our preconditioning strategies are highly effective for the problem considered.

Table 1: Number of interior point (Newton) iterations required to achieve convergence (blue, left), and average number of GMRES steps per interior point iteration before a relative preconditioned residual norm of 10^{-8} is achieved (black, right), for the test problem considered.

$\mathcal{P}_1$		$\beta=1$, $0\le y\le 1.5\times10^{-5}$, $0\le u\le 3\times10^{-4}$		$\beta=10^{-1}$, $0\le y\le 1.5\times10^{-4}$, $0\le u\le 3\times10^{-3}$		$\beta=10^{-2}$, $0\le y\le 1.5\times10^{-3}$, $0\le u\le 0.03$		$\beta=10^{-3}$, $0\le y\le 0.015$, $0\le u\le 0.3$		$\beta=10^{-4}$, $0\le y\le 0.08$, $0\le u\le 1.5$		$\beta=10^{-5}$, $0\le y\le 0.15$, $0\le u\le 5$	
h	2^{-2}	5	7.0	7	6.1	9	5.5	11	8.7	12	7.1	13	11.5
	2^{-3}	5	10.7	7	9.8	10	8.2	11	9.3	13	8.5	14	19.9
	2^{-4}	6	11.0	8	10.0	10	9.2	12	9.6	14	9.5	15	18.4
	2^{-5}	7	10.8	10	9.8	11	9.3	13	9.6	15	9.6	16	16.8
	2^{-6}	8	12.4	10	9.7	12	9.5	14	9.7	16	9.4	19	21.1
	2^{-7}	9	12.2	12	10.1	13	10.6	15	9.4	17	9.9	20	26.1

$\mathcal{P}_2$		$\beta=1$, $0\le y\le 1.5\times10^{-5}$, $0\le u\le 3\times10^{-4}$		$\beta=10^{-1}$, $0\le y\le 1.5\times10^{-4}$, $0\le u\le 3\times10^{-3}$		$\beta=10^{-2}$, $0\le y\le 1.5\times10^{-3}$, $0\le u\le 0.03$		$\beta=10^{-3}$, $0\le y\le 0.015$, $0\le u\le 0.3$		$\beta=10^{-4}$, $0\le y\le 0.08$, $0\le u\le 1.5$		$\beta=10^{-5}$, $0\le y\le 0.15$, $0\le u\le 5$	
h	2^{-2}	5	14.2	7	12.0	9	7.8	11	6.1	12	5.5	13	5.6
	2^{-3}	5	12.8	7	11.1	10	10.2	11	8.9	13	9.9	14	9.3
	2^{-4}	6	13.2	8	14.4	10	10.7	12	10.4	14	10.9	15	10.5
	2^{-5}	7	13.0	10	14.4	11	13.5	13	13.3	15	11.2	16	12.6
	2^{-6}	8	11.8	10	13.2	12	14.9	14	14.6	16	14.8	19	14.6
	2^{-7}	9	11.6	12	13.4	13	15.0	15	14.8	17	15.3	20	16.2

References

[1] M. Benzi, G. H. Golub, and J. Liesen, Numerical solution of saddle point problems, Acta Numerica **14**, pp. 1–137 (2005).
[2] J. W. Pearson and J. Gondzio, Fast interior point solution of quadratic programming problems arising from PDE-constrained optimization, submitted (2016).
[3] J. W. Pearson and A. J. Wathen, Fast iterative solvers for convection-diffusion control problems, Electronic Transactions on Numerical Analysis **40**, pp. 294–310 (2013).
[4] J. W. Pearson and A. J. Wathen, A new approximation of the Schur complement in preconditioners for PDE-constrained optimization, Numerical Linear Algebra with Applications **19**, pp. 816–829 (2012).